

Internetbefragung und deren Auswertung – leicht gemacht

DR. HELMUT STAUCHE, UNIVERSITÄT JENA

Mit Onlinebefragungen präsentiert sich gegenüber den traditionellen Befragungsmethoden eine Methode der Datenerhebung, die aus der aktuellen empirischen Praxis nicht mehr wegzudenken ist.

Onlinebefragungen werden oftmals für den Gebrauch im Internet vorbereitet; genauso gut können sie jedoch auch in Intranetzen, z.B. anstelle einer Gruppenbefragung mit Hilfe eines herkömmlichen Fragebogens, gute Dienste leisten.

Der für sie zu betreibende Aufwand ist – ein gutes Erstellungstool vorausgesetzt – geringer als im Falle der schriftlichen Befragung und die Kosten der Befragung liegen, verglichen mit der postalischen Versendung von Fragebögen, weitaus niedriger.

Es gibt mittlerweile keine triftigen Gründe mehr, die die Akzeptanz von Online-Befragungen schmälern könnten. Auch der Aspekt der Datensicherheit gilt heute als hinreichend erfüllt. Die Empiriker schätzen darüber hinaus, dass die Daten nicht in einem aufwändigen Verfahren von Hand in Rohdatentabellen überführt werden müssen, sondern direkt aus der Datenbank in das Statistikprogramm genommen werden können. Eine nicht zu unterschätzende Fehlerquelle wird hierdurch ausgeschaltet. Diese Prozedur ist bei Papierbogen-Befragungen nur dann vermeidbar, wenn eine – immer noch sehr teure – Anlage zum Kreieren und Scannen der ausgefüllten Fragebögen zur Verfügung steht.

Bezüglich der Repräsentativität der Daten gibt es bei Internetbefragungen – gemeint sind also nicht vom Forscher begleitete Erhebungen in organisierten Gruppen, wie Schulklassen – keine anderen Mängel als bei postalischem Versand von Fragebögen. Voraussetzung dafür ist jedoch, dass die Möglichkeit der Bearbeitung des Onlinefragebogens nicht vom technischen Standard des Computers des Befragten und von der Qualität seines Internetzuganges abhängt. Da heute diesbezüglich noch enorme Unterschiede existieren, empfiehlt sich der Einsatz einer schlichten Lösung, die für jeden Internetbrowser gleichermaßen tauglich ist und auch bei relativ niedrigen Übertragungsraten erfolgreich arbeitet.

Ist dies erfüllt, sehen wir keine spezifischen Nachteile der Onlinebefragung. Die existierenden unterscheiden sich nicht zwischen den beiden Medien, also zwischen dem Ausfüllen eines zugesandten Papierbogens und dem eines Bildschirmformulars, dessen URL dem Befragten per Email bekannt gegeben wurde. In beiden Fällen muss die Instruktion hieb- und stichfest sein, sie darf einerseits nicht langweilen, andererseits muss sie auf alle denkbaren auftretenden Fragen eine Antwort parat haben. Mängel dieser Art lassen sich durch einen vorgeschalteten Pilottest weitestgehend ausräumen.

Nicht zu lösen ist auch das Problem, dass es dem Empiriker unbekannt bleibt, wie zügig und unter welchen eventuell störenden Einflüssen die befragte Person arbeitet. Des Weiteren kann nicht ausgeschaltet werden, dass Familienangehörige oder andere während des Ausfüllens in der Nähe befindliche Personen einen Anteil am Antwortverhalten haben.

Auf dem Markt sind gegenwärtig viele und fast in jedem Falle kostenpflichtige Softwarelösungen zur Erstellung von Onlinefragebögen, zur Durchführung und in einigen Fällen auch zur Auswertung der Befragung. All diese erfüllen die Grundbedingung, dass ihre Anwender keinerlei Programmierkompetenz haben müssen. Insofern ist die Frage berechtigt, warum sich der Autor einer weiteren Lösung verschrieben hat. Zur Begründung gibt es gleich mehrere Argumente:

1. Das fertige Befragungsprogramm entsteht automatisch in PERL, einer Internetprogrammiersprache, die von allen WWW-Servern mit einer sogenannten CGI-Schnittstelle¹ verarbeitet wird.
2. Ein Perlscript ist ein reiner ASCII-Text, der Programmumfang ist demnach nicht zu unterbieten. Für eine Befragung mit einem Umfang von ca. 20 Fragen entsteht ein Script von ca. 10 kByte Größe, für dessen Öffnen selbst bei langsamem Netzzugang über ein analog arbeitendes Telefonmodem keine Wartezeit entsteht.
3. Die Eingabemethodik wurde so weit wie möglich vereinfacht: Auf einem einzigen Tabellenblatt vorbereitete Excel-Zellen sind mit Titel, Instruktion, Fragen

¹ Dies ist Standard.

und – im Falle von Single und Multiple Choice² – Auswahlantworten zu belegen. Dies ist für die Herstellung weitaus angenehmer und effektiver als die übliche sukzessive Abfrage von Fragen und dazu gehörigen Auswahlantworten.

4. Die entstehende Datenbank³ ist ebenfalls eine reine Textdatei, womit kurze Zeiten für das Absenden der Daten garantiert sind. Die Daten jeder befragten Person erscheinen als eine Textzeile, die Variablen – d.h. die freien Antworten, die Antwortcodes, das Datum und der bei nicht anonymer Befragung einzugebende Name – werden durch das Zeichen | getrennt.
5. Sollte der Benutzer nur „normale“ Ansprüche an die Auswertung haben, wird ihm fast die gesamte Arbeit abgenommen. Unter diesen Ansprüchen sind einfache Häufigkeitsanalysen mit absoluter und relativer Ausgabe, die Ausgabe der Deskriptiva metrischer (und quasimetrischer) Variablen, die Herstellung gebündelter Variablen aus Einzelvariablen gleicher Kodierung⁴ sowie drei inferenzstatistische Analysen (Kreuztabellenanalyse mit χ^2 , Pearson'sche Korrelationsanalyse und unabhängiger t-Test zum Prüfen von Mittelwertunterschieden)⁵ zu verstehen.

Darüber hinaus können alle Ergebnisse mit den zugehörigen Diagrammen kopier- bzw. ausdrucksfähig abgerufen werden.

Jedes Perlscript für den Zweck der Onlinebefragung hat eine invariable Struktur. Einige Teile des Scripts liegen fest, die anderen modifizieren sich in Abhängigkeit von Anzahl, Kategorie und Inhalt der Fragen.

² Als Frage-/Eingabetypen stehen zur Verfügung:

- Single Choice für ordinale oder nominale Variablen
- Single Choice für quasimetrische Variablen
- Single Choice in Tabellenform (Kopfzeile mit den Wertelabels, dazu beliebig viele Einzelfragen in den folgenden Tabellenzeilen)
- Multiple Choice
- freier Text mit festlegbarer maximaler Zeichenanzahl
- Zahl (Ganzzahl oder Dezimalzahl) mit festlegbarer Stellenzahl und
- Kommentar (für Zwischenüberschriften, für weitere Instruktionen, zur Einleitung von Fragen der Typen Multiple Choice und Single Choice in Tabellenform)

³ Mit diesem Begriff ist hier Beitrag keine Datenbank im Sinne der EDV gemeint, sondern lediglich die Datei, in die die Antworten geschrieben werden.

⁴ siehe S. 9

⁵ siehe S. 10

Für die Herstellung des Programms wurde deshalb ein einfacher wie genialer Weg beschritten. Ein in VBA⁶ für Excel geschriebenes Programm, das nach der oben beschriebenen inhaltlichen Vorarbeit mit einem Klick ausgelöst wird, übernimmt die Perlscriptherstellung. Danach bleibt nur noch das Hochladen des fertigen Scripts (evtl. auch der leeren als Datenbank fungierenden Textdatei⁷) auf den WWW-Server und das Setzen der Ausführungs- und Schreibrechte.

Zu all dem muss der Anwender von Perlprogrammierung überhaupt nichts verstehen und es ist völlig belanglos, wenn er die entstandene Scriptsyntax nicht versteht.

Versierte Anwender können auf die Lektüre der nachfolgenden Anleitung zum Umgang mit dem Perlscriptgenerator verzichten, weil in die betreffenden Excel-Zellen per Zellenkommentar alle wesentlichen Hinweise in Kurzform aufgenommen wurden.

Im Anhang 1 sieht man den Screenshot – und damit nur einen Ausschnitt – des Excelblattes **herstellung** mit einem einblendbaren Beispiel. Anhang 2 zeigt das Browserbild zum selben Beispiel, so wie es der Befragte sieht.

Im Jahre 2010 wurde diese Arbeit ergänzt: Ausgangspunkt für die Ergänzung war die Notwendigkeit, dass im Rahmen des Praxissemesters im Lehramtstudium nach dem Jenaer Modell jede(r) Studierende ein überschaubares, schulbezogenes empirisches Forschungsprojekt zu absolvieren hat. Da auf eine empirisch-statistische Vorbildung nicht aufgebaut und diese im Begleitseminar nur in Ansätzen geboten werden kann, entstand der Gedanke, die reichhaltigen Auswertungsfeatures auch für eine schriftliche Befragung zur Verfügung zu stellen.⁸

Nachwort

Die beschriebene VBA-Programmierung leistet das Angekündigte – nicht mehr und nicht weniger. Einer Anwendung der Software für Ihre Zwecke steht nichts im Wege. Nehmen Sie zu diesem Zweck bitte mit mir Kontakt auf.

⁶ Visual Basic for Applications

⁷ unnötig bei Verwendung von AFS (Andrew File System)

⁸ siehe S. 12

Anleitung zum Perlscriptgenerator für die Internetbefragung sowie zur anschließenden Auswertung des Datensatzes

programmiert in VBA für Excel von Dr. H. Stauche, 2004+

Sie müssen von Perl nichts verstehen, um ein Perlscript herzustellen, das auf einem Internetserver laufen kann und eine Onlinebefragung ermöglicht. Möglich ist die Verwendung von Single Choice, Multiple Choice, freier Texteingabe, Zahleingabe, und Eingabe von Kommentaren (für Zwischenüberschriften, weitere spezielle Instruktionen, Einleitung von Fragen etc.) in beliebiger Reihenfolge⁹ und Häufigkeit.

Dabei kann Single Choice als einfache Frage mit ihren Auswahlantworten oder ein gesamte Paket von Fragen mit untereinander gleichen Auswahlantworten gewählt werden: Bei der Wahl der letzteren Option werden Fragen und Auswahlantworten in der Internetansicht in Tabellenform dargestellt.

Des Weiteren besteht die Möglichkeit der Gestaltung spezieller Single-Choice-Fragen mit quasimetrischen¹⁰ Werten. Das heißt, dass in der späteren Auswertung diese Variablen wie metrische behandelt werden: Mit Hilfe des projektinternen Auswertungsmoduls können Deskriptiva, Korrelationen und der unabhängige t-Test berechnet werden. Die Labels einer solchen Variablen könnten – um Äquidistanz zu gewährleisten – wie in den folgenden Beispielen aussehen¹¹:

Beispiel „Skalenwerte“

0	1	2	3
---	---	---	---

Beispiel „Sprachsymmetrie“

trifft nicht zu	trifft eher nicht zu	trifft eher zu	trifft zu
-----------------	----------------------	----------------	-----------

Beispiel „Symbolik“

--	-	±0	+	++
----	---	----	---	----

Anders als bei der automatischen Wertevergabe bei den anderen Single-Choice-Variablen, deren kleinster Wert die Eins ist, wird bei den quasimetrischen mit der Null begonnen. Der Range dieser Variablen reicht demnach von Null bis zur um eins verringerten Werteanzahl.

Für die Bearbeitung des Fragebogens kann zwischen den Versionen *anonym* oder *nicht anonym* umgeschaltet werden.

⁹ Bezüglich der Reihenfolge sollte aber beachtet werden, dass die unabhängigen Variablen unter den abhängigen stehen. Werte unabhängiger Variablen sind zum Beispiel „männlich – weiblich“ oder „Stadt – Land“ oder „ohne Partner – mit Partner“. Andernfalls entstehen ungünstige Kreuztabellen und daraus berechnete Diagramme.

¹⁰ Als quasimetrisch darf man eine eigentliche ordinale Variable dann auffassen – und mit ihr dann auch rechnen wie mit einer echten metrischen –, wenn deren Werte Symmetrie aufweisen.

¹¹ Die in schriftlichen Papierfragebögen vorzufindende – und dort durchaus probate – Möglichkeit, nur die Extremwerte zu attribuieren, ist in diesem Programm nicht möglich, weil die Zellen aller Werte besetzt werden müssen.

Vorbereitungen

1. Legen Sie auf Ihrem CGI-fähigen¹² Server ein Verzeichnis für Ihre Befragung an.

Arbeit mit dem Programm

1. Öffnen Sie in MS Excel die Datei **befragung.xls** und aktivieren Sie die Makros¹³.
2. Lesen Sie dazu zunächst die Kommentare, indem Sie die Maus auf die Zellen mit dem kleinen roten Dreieck führen. Verfahren Sie genau nach den Anweisungen und kontrollieren Sie die Richtigkeit. Sie haben die Möglichkeit zum Schreiben oder Einfügen einer Instruktion. Berücksichtigen Sie dazu die üblichen Standards.
3. Sie werden den Befragungsinhalt sicherlich in einem Textprogramm vorbereitet haben und müssen in diesem Falle fast nichts von Hand eintippen. Öffnen Sie sowohl Ihren Entwurf im Textprogramm und auch **befragung.xls**. Markieren Sie im Textprogramm jeweils einen Teil aus dem Entwurf und bringen Sie diesen mit der hellgrünen Schaltfläche **TEXT AUS EINEM TEXTPROGRAMM OHNE ZEILENUMBRUCH IN DIE AKTIVE ZELLE EINFÜGEN**¹⁴ in die entsprechende Excel-Zelle. Es sollte prinzipiell mit dieser hellgrünen Schaltfläche gearbeitet werden. Normales Kopieren–Einfügen würde Zeilenumbrüche mitnehmen und Text, der für eine Zelle vorgesehen ist, in mehrere Excel-Zellen kopieren.
4. Versäumen Sie nicht – auch wenn Ihre Befragung anonym ablaufen soll – demografische Variablen aufzunehmen. Meist sind dies Single-Choice-Fragen.
5. Nach der Kontrolle, die Sie mit dem Button **PERLSCRIPT HERSTELLEN** abschließen, schaltet das Programm in das Tabellenblatt **programm** um.
6. In diesem sehen Sie ab Zeile 3 das entstandene Perlscript. Klicken Sie auf den Button **SPEICHERN**. Sie werden nunmehr gefragt, ob Ihre Datenbank denselben Namen vor der Dateierweiterung **.txt** haben soll wie Ihr Perlscript. Dieses Feature wurde aus Datenschutzgründen eingearbeitet. Heißt zum Beispiel Ihr Perlscript **maerchen.pl**, würde ohne diese Abfrage Ihre Datenbank automatisch mit dem Namen **maerchen.txt** angelegt. Dies eröffnet versierten Nutzern die Möglichkeit, den Namen auszudenken und Einblick in die Datenbank zu nehmen. Ein völlig frei gewählter, nicht erratbarer Name macht einen Einblick nahezu unmöglich.
Das Perlscript wird nach dieser Entscheidung in den Zwischenspeicher kopiert und ein Editorfenster (Fenster von Windows Notepad) öffnet sich. Mit **Strg+v** (oder: Bearbeiten => Einfügen) kopieren Sie das Script in den Editor. Speichern Sie es dann unter dem von Ihnen vergebenen Namen mit der Erweiterung **.pl**, entweder gleich auf dem Server oder zunächst auf Ihrem PC

¹² für die Verwendung des WWW-Servers der FSU Jena erfüllt, aber auch sonst allgemeiner Standard

¹³ Das sind die hinter den Schaltflächen (Buttons) in VBA (Visual Basic für Applications) geschriebenen Programme, die zum Erfüllen der Aufgaben nötig sind. Dass ihre Aktivierung abgefragt wird, ist eine reine Sicherheitsvorkehrung.

¹⁴ Natürlich funktioniert auch Bearbeiten/Einfügen (Strg+v), wenn es sich um eine Textmarkierung ohne festen Zeilenumbruch handelt. Liegt Letzteres vor, dann muss die Schaltfläche verwendet werden, weil andernfalls der Text in mehrere Excel-Zellen geschrieben wird.

oder einem Datenträger. Wenn Sie beispielsweise den Programmnamen **maerchen** vergeben haben, lautet der Dateiname **maerchen.pl**.

7. Anschließend stellen Sie mit dem **LÖSCHEN**-Button den Ausgangszustand wieder her. Dabei werden sie Sie gefragt, ob Sie die Datei **befragung.xls** unter Ihrem vergebenen Namen speichern wollen. Antworten Sie mit **Ja**, wenn Sie die Auswertung des entstandenen Datensatzes mit Hilfe desselben Programms auswerten möchten. Für das erwähnte Beispiel würde die Datei **befragung.xls** automatisch unter dem Namen **maerchen.xls** gespeichert werden.
8. Das Script kann im Internet erst dann funktionieren, nachdem...
 - es sich im Serververzeichnis befindet
 - dort eine leere als Datenbank dienende Textdatei mit demselben Namen (für das genannte Beispiel hieße diese **maerchen.txt**) existiert¹⁵ und
 - die Rechte beider Dateien (des Perlscripts und der Textdatei) auf dem Server auf 777 (Lesen, Schreiben und Ausführen für alle) gesetzt wurden.¹⁶
9. In die Datenbank werden am Ende jeder Zeile das Datum und die Uhrzeit des Absendens im Format **ddmmyyyyhhmm** mitgespeichert.

Um die Datenbank auszuwerten, gibt es zwei Möglichkeiten:

Erster Fall: Für die Auswertung genügen einfache Analysen.

Dann kann einfach mit dem Programm weiter gearbeitet werden, das nachfolgend erläutert wird:

1. Allgemeines und Häufigkeiten

- 1.1 Wenn Sie nach dem Löschen des Perl-Programms gefragt werden, ob Sie die Datei **befragung.xls** mit dem konkreten Inhalt Ihrer Befragung speichern wollen, um später mit dieser die Auswertung vornehmen zu können, dann antworten sie **Ja**. Es entsteht dann automatisch der neue Dateiname (in meinem Beispiel **maerchen.xls**) und die Datei wird im selben Verzeichnis abgelegt, in dem sich auch die schreibgeschützte Datei **befragung.xls** befindet. In dieser Datei sehen Sie zwei nunmehr zwei bis dahin verborgene Tabellenblätter **auswertung** und **datenbank**, die Sie später benötigen.
- 1.2 Nun lassen Sie die Befragung bis zum Endzeitpunkt laufen.
- 1.3 Laden danach die Datenbank (im Beispiel **maerchen.txt**) vom Server herunter in dasselbe Festplattenverzeichnis, in dem sich auch Ihre spezielle Datei befindet (im Beispiel **maerchen.xls**).

¹⁵ Diese Datei kann man erzeugen, indem man unter Start → Programme → Zubehör den Editor aufruft und eine leere Datei unter dem oben angegebenen Namen speichert.

Bei Verwendung des Andrew File System (AFS) kann unter Zugriffssteuerlisten → WWW-Server → i (für insert) angehakt werden, dann ist das Anlegen der Textdatei automatisiert.

¹⁶ Bei Verwendung des Andrew File System (AFS) werden diese Rechtecodes nicht ausgewertet. Dafür ist aber das Schreibrecht für das verwendete Serververzeichnis zu setzen: Zugriffssteuerlisten → WWW-Server → w (für write) anhängen.

1.4 Gehen Sie auf das Tabellenblatt **auswertung** und klicken Sie **DATEN EINLESEN**.

Es öffnet sich die o.g. Textdatei (im Beispiel **maerchen.txt**) und deren Daten werden eins zu eins in das Blatt **datenbank** übertragen. In die Kopfzeile werden die Fragen geschrieben, eine Kontrollmöglichkeit für richtiges Einlesen.

1.5 Klicken Sie nun den Button **HÄUFIGKEITSANALYSE**.

Es erscheint eine Tabelle mit allen Fragen, deren (Auswahl-)antworten und Angaben zur absoluten und relativen (prozentualen) Häufigkeit.

1.6 Um einfache Diagramme zu den Häufigkeitstabellen zu erzeugen, doppelklickt man auf eine oder mehrere rosafarbene Zellen in den Spalten 7 und/oder 8 (die Zelle wird nach dem Klicken grün) und danach auf den Button **HÄUFIGKEITSDIAGRAMME**. Die Auswahl in Spalte 7 steht für ein Balkendiagramm der absoluten Häufigkeit, in Spalte 8 dagegen für ein Balkendiagramm der prozentualen Häufigkeit.

Es ist möglich, für einen Teil der Variablen absolute und für einen anderen Teil prozentuale Balkendiagramme gleichzeitig herstellen zu lassen.

1.7 Mit dem Button **BLATT AUSWERTUNG DRUCKEN** wird der reine Ergebnisteil des Blattes **auswertung** – also ohne die die Buttons enthaltende erste Zeile – im Querformat gedruckt.

1.8 Durch das Klicken des Buttons **AUSWERTUNG ZU EINZELNER FRAGE DRUCKEN** reduziert man den Ausdruck auf die ausgewählte Frage.

1.9 Schließlich steht noch der Button **AUSWERTUNG ZU EINZELNER FRAGE IN WORD** zur Verfügung. Damit werden dieselben Zeilen wie unter 1.8 in den Zwischenspeicher kopiert, so dass sie in ein qualifiziertes Textprogramm¹⁷ (außer in MS Word auch in OpenOffice) eingefügt werden können.

2. Gebündelte Variablen

In vielen Anwendungsfällen ist es sinnvoll, untereinander gleich kodierte Variablen zu einer komplexeren zusammenzufassen – zu bündeln. Das Auswertungsprogramm sieht diese Möglichkeit ebenfalls vor. Folgendermaßen ist vorzugehen:

In der Spalte 6 sind diejenigen gelben Felder doppelzuklicken (mindestens zwei), deren Variablen gebündelt werden sollen. Beachten Sie dabei, dass die Kodierung dieser Variablen untereinander gleich sein muss.

Für die Berechnung der neuen Variablen stehen 4 Modi zur Verfügung, wobei einer gewählt werden muss:

1 = Der geringste Wert der Variablen steht oben. Die Skala soll bei Null beginnen.

2 = Der geringste Wert der Variablen steht oben. Die Skala soll bei Eins beginnen.

3 = Der geringste Wert der Variablen steht unten. Die Skala soll bei Null beginnen.

4 = Der geringste Wert der Variablen steht unten. Die Skala soll bei Eins beginnen.¹⁸

Die neue gebündelte Variable im Blatt **auswertung** unter der jeweils letzten Variablen angezeigt. Berechnet werden das arithmetische Mittel, die Varianz und die Standardabweichung.

¹⁷ Editoren oder Wordpad sind ungeeignet.

¹⁸ „oben“ und „unten“ beziehen sich auf die Position des Variablenwertes im Blatt **auswertung**. Die ‚geringste‘ Ausprägung ist bei leistungsbeschreibenden Variablen die schlechteste.

Des Weiteren wird die neue Variable in der Spalte 5 (Spaltenüberschrift *metrisch o. quasimetrisch*) gelb markiert, so dass mit ihr sowohl die Korrelationsanalyse als auch der unabhängige t-Test wie mit einer herkömmlichen metrischen Variablen möglich wird.

3. Deskriptive Statistik

3.1. Deskriptiva metrischer Variablen

Um Deskriptiva metrischer (z.B. das Alter der Befragten in Jahren) oder auch quasimetrischer Variablen zu berechnen, doppelklicken Sie in der Spalte 5 des Blattes **auswertung** (Spaltenüberschrift *metrisch o. quasimetrisch*) auf eine oder mehrere gelbe Zellen (die Zelle wird nach dem Klicken grün) und bedienen Sie danach die Schaltfläche **DESKRIPTIVA**. Das Programm schaltet in das gleichnamige Tabellenblatt um und zeigt Ihnen die deskriptiven Statistiken Minimum, Maximum, arithmetisches Mittel, Varianz, Standardabweichung, mittlere Abweichung, Quartilsabstand, Median und Modus¹⁹ für die ausgewählte(n) Variable(n).²⁰

Wenn Sie mehrere Variablen ausgewählt haben, dann werden die Ergebniszeilen der einzelnen Variablen alternierend mit zwei Farben unterlegt.

3.2. Kreuztabellenanalyse

Doppelklicken Sie im Blatt **auswertung** auf zwei gelbe Zellen in Spalte 6 (Spaltenüberschrift *ordinal/nominal*, aber auch *quasimetrisch* wird markiert). Dabei ist die Reihenfolge nur dann beliebig, wenn es sich um zwei gleichberechtigte Variablen handelt, also solche, bei denen es keine abhängige und unabhängige gibt. Ist dies nicht der Fall, ist also z.B. die Variable A determinierend für die Variable B, dann bezeichnet man B als (von A) abhängige Variable. Dann muss zuerst A – die unabhängige Variable – doppelgeklickt werden und danach B – die abhängige. Auch deren Zellen werden nach dem Klicken grün und zeigen mit 1 und 2 die Reihenfolge des Anklickens an.

Bedienen Sie nun die Schaltfläche **KREUZTABELLE**. Das Programm schaltet in das gleichnamige Tabellenblatt um und zeigt Ihnen die fertige Kreuztabelle für die beiden ausgewählten Variablen.

Kreuztabellen sind also herstellbar für eine beliebige Kombination aus Single-Choice- und Multiple-Choice-Variablen.

Weitere Funktionen auf diesem Tabellenblatt sind:

- **Kreuztabelle ausdrucken** (Nur die reine Tabelle wird gedruckt, nicht der Rest des Blattes.)
- **Diagramm erzeugen** (Es entsteht ein nach der unabhängigen Variablen gruppiertes Balkendiagramm.)

¹⁹ Die Berechnung der Modi erfolgt so, dass beginnend mit Spalte 10 alle Modi angezeigt werden. Die Berechnung richtet sich nach der Definition, dass ein Modus in der geordneten Folge der Messwerte einen weniger häufigen linken und rechten Nachbarwert haben muss. Sind benachbarte Werte gleich häufig vertreten und haben als Paar einen weniger häufigen linken und rechten Nachbarwert, dann spricht man von einem breitgipfligen Modus. Die Attribute einfacher bzw. breitgipfliger Modus werden mitgeliefert.

²⁰ Werden auch Dezimalzahlen als Werte erkannt, wird die Berechnung des Medians und der Modi nicht vorgenommen, weil diese Deskriptiva nur für Ganzzahlen sinnvoll sind.

4. Inferenzstatistik

4.1. χ^2 -Test

Die Kreuztabelle lässt sich prüfstatistisch mit Hilfe des χ^2 -Zusammenhangstests weiter auswerten. Beachten Sie, dass für diesen Test nur zwei Single-Choice-Variablen vom Variablentyp s oder sz oder sq in Frage kommen. Eine Kombination einer Multiple-Choice-Variablen mit einer Single-Choice-Variablen oder ganz und gar die Kombination zweier Multiple-Choice-Variablen darf nicht getestet werden).

Der χ^2 -Test wird durchgeführt, indem Sie im Blatt **kreuztabelle** auf den gleichnamigen Button **chi²-test durchführen** klicken. Angezeigt werden der χ^2 -Wert sowie der Ausgang der Signifikanzprüfung: *hochsignifikanter Zusammenhang* [1%-Niveau], *signifikanter Zusammenhang* [5%-Niveau] oder *nicht signifikanter Zusammenhang*.)

Für zwei Fälle gibt es Sonderausgaben:

- Alle Zeilen- und Spaltensummen sind zwar > 0 , aber es gibt mindestens eine Zelle, die mit weniger als 5 Häufigkeiten besetzt ist. Dann wird ausgegeben, dass die Signifikanzprüfung nur mit gebotener Vorsicht interpretiert werden darf.
- Mindestens eine Zelle ist leer. Dann wird χ^2 nicht berechnet und das Nichtberechnen damit begründet.

4.2. Korrelationsanalyse

Doppelklicken Sie im Blatt **auswertung** auf zwei gelbe Zellen in der Spalte 5 (Spaltenüberschrift *metrisch o. quasimetrisch*), also auf rein metrische oder auch auf quasimetrische Variablen (die Zellen werden nach dem Klicken grün), und bedienen Sie danach die Schaltfläche **KORRELATION**. Das Programm schaltet in das gleichnamige Tabellenblatt um und zeigt Ihnen als Ergebnis den Pearson'schen Maßkorrelationskoeffizienten für die beiden ausgewählten Variablen unter Angabe der Signifikanz.²¹

Zur Korrelationsanalyse können auch neu hergestellte, gebündelte Variablen herangezogen werden.

4.3. Unabhängiger t-Test zur Prüfung von Mittelwertunterschieden

Doppelklicken Sie im Blatt **auswertung** eine gelbe Zelle in der Spalte 5 (Spaltenüberschrift *metrisch o. quasimetrisch*) – also auf die metrische oder auch quasimetrische Untersuchungsvariable – und zur Auswahl der Gruppierungsvariablen eine gelbe Zelle in der Spalte 6 (Spaltenüberschrift *ordinal/nominal*). Achtung: Die Gruppierungsvariable muss unbedingt dichotom (d.h. zweiwertig) sein.

Bedienen Sie danach die Schaltfläche **UNABHÄNGIGER T-TEST**. Das Programm schaltet in das gleichnamige Tabellenblatt um und zeigt Ihnen das Testergebnis mit folgenden Details:

- Namen der untersuchten Variablen und der gruppierenden Variablen

²¹ Die kritischen Vergleichswerte zur Signifikanzprüfung wurden Claus/Ebner (1974) entnommen: Tafel 9, Zahl der Freiheitsgrade und Zufallshöchstwerte des Korrelationskoeffizienten (Claus, G. & Ebner, H.: Grundlagen der Statistik. Volk und Wissen Volkseigener Verlag Berlin 1974, S. 492)

- arithmetisches Mittel der untersuchten Variablen für den Wert 1 und für den Wert 2 der gruppierenden Variablen
- Standardabweichungen der untersuchten Variablen für den Wert 1 und für den Werte 2 der gruppierenden Variablen²²
- Wert der Prüfgröße t
- Ergebnis der Signifikanzprüfung: *hochsignifikanter Unterschied* [1%-Niveau], *signifikanter Unterschied* [5%-Niveau] oder *nicht signifikanter Unterschied*.²³

Beim unabhängigen t-Test kann auch eine gebündelte Variable als Untersuchungsvariable dienen.

5. Gruppieren von Textantworten

Auch für die frei zu beantwortenden Fragen vom Typ **t** über die bloße Auflistung hinaus gehende Auswertungsmöglichkeit zur Verfügung.

Wenn Sie in der Spalte 6 (Spaltenüberschrift *ordinal/nominal*), in der alle Fragen vom Typ s, sz, sq, m und t gelb markiert sind, eine Textfrage (t) und eine Single-Choice-Frage vom Typ (s) oder auch (sq) grün klicken, dann wird nach dem Ausführen des Buttons **TEXTANTWORTEN GRUPPIEREN** eine nach den Werten der Single-Choice-Frage gesplittete Liste der Textantworten ausgegeben.

6. Fehler

Sollten Sie in Vorbereitung der Prozeduren Fehler gemacht haben (z.B. falscher Datentyp, falsche Werteanzahl, ungleiche Kodierung innerhalb desselben Datentyps für die Bündelung, keine, zu viel oder zu wenig Variablen), dann endet der Programmablauf zunächst und Sie werden zur Korrektur aufgefordert, um danach neu zu beginnen.

7. Löschfunktionen

Das Löschen kann mit Hilfe von vier Buttons differenziert erfolgen:

- **Alle Markierungen löschen** ist dann sinnvoll, um eine neue Prozedur vorzubereiten, besonders dann, wenn nicht alle Markierungen der vorausgegangenen Arbeit ohne zu Scrollen auf dem Bildschirm zu sehen sind.
- **Alle Diagramme löschen** ist sinnvoll, um die Anzeige der geöffneten Blätter in der Fußleiste von Excel nicht überzustrapazieren. Günstig ist es, mehrere Diagramme in Serie herzustellen, diese in MS Word zu kopieren (normales Kopieren-Einfügen-Verfahren) und danach in MS Excel zu löschen.
- **Alle Ergebnisse löschen** sowie **Alle Daten und Ergebnisse löschen** bedürfen keiner Erläuterung.

Vor allen Löschvorgängen gibt es eine Sicherheitsabfrage.

²² Aus beiden Standardabweichungen wird im Hintergrund deren sog. gewogenes Mittel berechnet, das in die Formel für t einfließt.

²³ Die kritischen Vergleichswerte zur Signifikanzprüfung wurden Claus/Ebner (1974) entnommen: Tafel 3, t-Verteilung (Claus, G. & Ebner, H.: Grundlagen der Statistik. Volk und Wissen Volkseigener Verlag Berlin 1974, S. 478)

Zweiter Fall: Weitere Prozeduren sollen möglich sein.

In diesem Falle sollte nach Abschluss der Befragung die Datenbank (im Beispiel: **maerchen.txt**) in ein Computerstatistikprogramm eingelesen werden.

Falls SPSS eingesetzt wird:

1. Zum Einlesen der Daten ist die Funktion *Datei → Textdaten lesen* zu verwenden. Als Trennzeichen ist unter „Andere“ das Zeichen | anzugeben (zu erzeugen mit AltGr + <).
2. Die Variablen erscheinen in der Reihenfolge des Fragebogens.
3. Die Codes der „normalen“ Single-Choice-Fragen lauten 1 bis n in der Reihenfolge des Fragebogens – von links nach rechts gezählt.
4. Die Codes der quasimetrischen Single-Choice-Fragen lauten 0 bis n-1 in der Reihenfolge des Fragebogens – von links nach rechts gezählt.
5. Bei Multiple-Choice-Fragen ist jede Auswahlantwort eine Variable, die entweder als 1 (= angekreuzt) oder als Fehlwert (= nicht angekreuzt) codiert ist.
6. Freie Texteingaben erscheinen automatisch im String-Variablenformat.
7. Nach dieser Zuordnung sind die Variablen- und Wertelabels zu vergeben und der Datensatz anschließend auszuwerten.
8. Die Auswertung des Fragetyps Multiple Choice erfolgt in SPSS mit Hilfe von *Analysieren → Mehrfachantworten*.

Dritter Fall: Die Auswertungsprozeduren des Programms befragung.xls sollen für eine schriftliche Befragung genutzt werden.

Zunächst ist im Blatt **herstellen** alles – also der Programmname, der Titel, die Instruktion, die Fragen mit deren Auswahlantworten bzw. mit den Angaben zur Antwortlänge bei Text und Zahl, und vor allem die richtige Kategorie der Fragen – einzutragen. Bis hierhin unterscheidet sich das Vorgehen nicht vom oben beschriebenen (vgl. S. 6, Punkte 1 bis 4).

Anstelle der Perlscriptherstellung klickt man auf die Schaltfläche **NUR AUSWERTEN**. Das Blatt **datenbank** öffnet sich mit ausgefüllter Kopfzeile. Dort stehen die Fragen sowie im Zellenkommentar die für die jeweilige Variablenkategorie gültige Kodierungsvorschrift. Nachdem die Antworten in den Fragebögen in kodierter Form in die Datenbank ab Zeile 2 übertragen wurden, kann die Auswertung wie unter Erster Fall: ...²⁴ erfolgen.

²⁴ siehe S. 7ff.

Microsoft Excel - maerchen.xls

Frage hier eingeben

10 F K U €

Anal

Bearbeitung zurücksenden... Bearbeitung beenden...

ZS2 maerchen

	1	2	3	4	5	6	7
1	Programme zur Herstellung einer Internetbefragung und deren anschließenden Auswertung Dr. H. Stauche 2004+						
2	Name	maerchen					
3	Titel	Was wissen Sie über Märchen?					
4	Instruktion	Füllen Sie den Fragebogen aus!					
5	Nummer	Formulierung der Frage	Kat	Kurzname			
6	1	Welche Märchen können Sie aus dem Internet finden?	m	var1	Dornröschen	Aschenputtel	Das tapfere Schneiderlein
7	2	Nennen Sie ein Märchen, das Sie für das Internet geeignet finden.	t	var2		30	
8	3	Danach kommen noch zwei Fragen.	k	var3			
9	4	Welches der nachstehenden Märchen kennen Sie?	s	var4	Der Wolf und die 7 Geißlein	Rapunzel	Rotkäppchen
10	5	Bewerten Sie den erwarteten Erziehungsprozess.	sk	var5	sehr gut	gut	befriedigend
11	6	Der Wolf und die 7 Geißlein	sz	var6			
12	7	Rotkäppchen	sz	var7			
13	8	Aschenputtel	sz	var8			
14	9	Die goldene Gans	sz	var9			
15	10	Das tapfere Schneiderlein	sz	var10			
16	11	Warum lesen Sie Kindern gern Märchen vor?	m	var11	Sie können aus den Märchen viel lernen.	Sie machen dann so große Augen.	Sie schlafen schneller ein.
17	12	Schätzen Sie allgemein die Erziehungsqualität.	sq	var12	sehr niedrig	niedrig	mittel
18	13	Ihr Geschlecht?	s	var13	männlich	weiblich	
19	14	Ihr Alter?	z	var14		3	
20	15	Danke für Ihre Mitarbeit.	k	var15			
21							
22							

Anleitung:
Zuerst Kommentare lesen (Zellen mit rotem Dreieck), dann weiße Zellen ausfüllen.
Am Schluss die Schaltfläche PERL-PROGRAMM HERSTELLEN ausführen.
Mit den dunkelblauen Buttons kann ein Beispiel ein- und ausgeblendet werden.

Text aus einem Textprogramm ohne Zeilenumbbruch in die aktive Zelle einfügen

BEISPIEL WEG **BEISPIEL ZEIGEN** **PERLSCRIPT HERSTELLEN** **NUR AUSWERTEN**

Anmerkung:

Wegen der Spaltenbreiten sind im Screenshot nicht alle Beispielfragen/-kommentare voll sichtbar. Entnehmen Sie den vollständigen Wortlaut der nächsten Grafik.

Befragung (© des Scripts by Dr. H. Stauche, 2010) - Mozilla Firefox

Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

http://www2.uni-jena.de/erzwiss/maerchen.pl

IBK IB CMS LSF Akt Mecker Begleit Meth TeamSpace Webmail SP CC CB

Befragung (© des Scripts by Dr. H. Stauche, 2010)

Was wissen Sie über Märchen?

Füllen Sie den Fragebogen aus!

Jetzt kommen die Fragen:

Welche Märchen können Sie aus dem Kopf erzählen? (Mehrfachbeantwortung ist möglich!)

☐ Dornröschen
☐ Aschenputtel
☐ Das tapfere Schneiderlein
☐ Hans im Glück

Nennen Sie ein Märchen, das Sie für besonders grausam halten. (maximal 30 Zeichen möglich!)

Danach kommen noch vier Fragen. Halten Sie durch!

Welches der nachstehenden Märchen können Sie am besten erzählen?

☐ Der Wolf und die 7 Geißlein
☐ Rapunzel
☐ Rotkäppchen
☐ Aschenputtel

Bewerten Sie den erwarteten Erziehungserfolg durch die nachfolgenden Märchen.

	sehr gut	gut	befriedigend	ausreichend	mangelhaft	ungenügend
Der Wolf und die 7 Geißlein	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rotkäppchen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aschenputtel	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die goldene Gans	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das tapfere Schneiderlein	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Warum lesen Sie Kindern gern Märchen vor? (Mehrfachbeantwortung ist möglich!)

☐ Sie können aus den Märchen viel lernen.
☐ Sie machen dann so große Augen.
☐ Sie schlafen schneller ein.

Schätzen Sie allgemein die Erziehungskompetenz von Märchen ein.

☐ sehr niedrig
☐ niedrig
☐ mittel
☐ hoch
☐ sehr hoch

Ihr Geschlecht?

☐ männlich
☐ weiblich

Ihr Alter? (maximal 1 Kommastelle möglich!)

Danke für Ihre Mitarbeit.

Fertig

Start SpeedProject's Speed... Befragung (© des Sc... XnView [gb]